ETH zürich

Modeling Load Forecast Uncertainty Using Generative Adversarial Networks

Yi Wang*, Gabriela Hug*, Zijie Liu#, Ning Zhang#

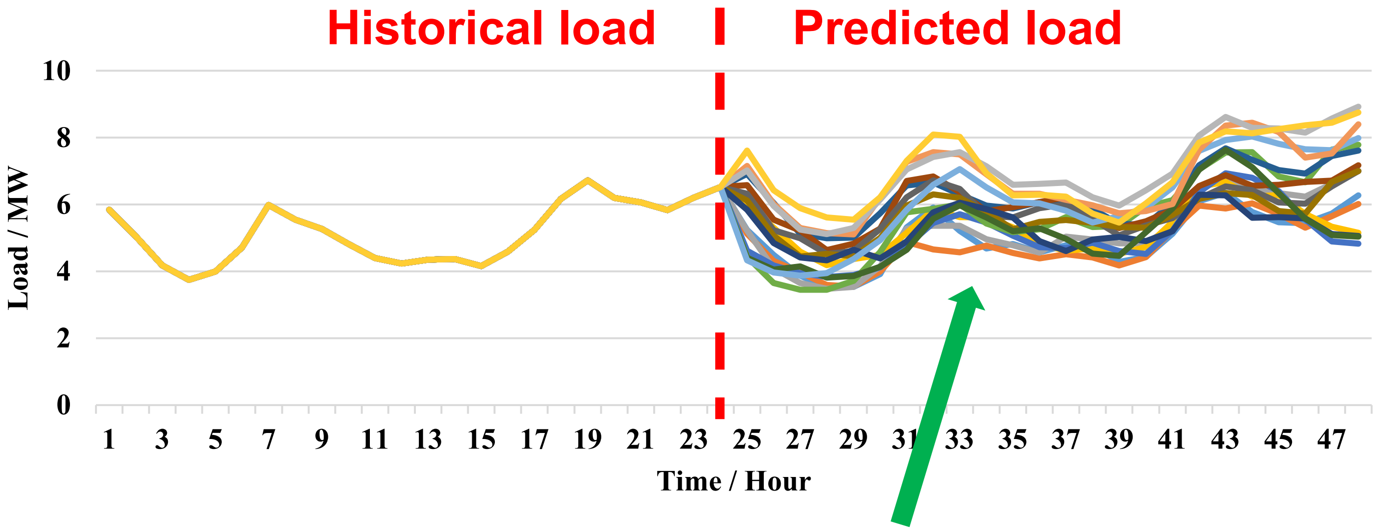
* Power Systems Laboratory, ETH Zurich

Department of Electrical Engineering, Tsinghua University

PSCC2020 | 3rd July, 2020 | Porto, Portugal

Contents

Introduction



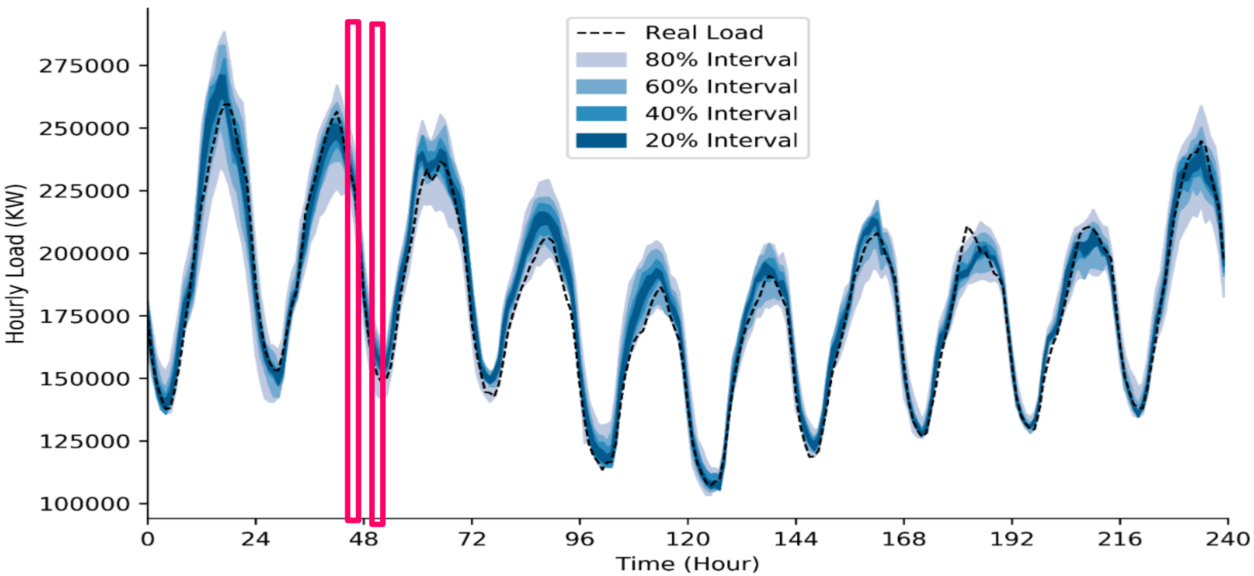
Compared with deterministic forecasting, probabilistic load forecasts provide comprehensive information about future uncertainties.

Current works:

- **Modeling uncertainties of the input features:** temperature scenarios generation, bootstrap sampling of features;
- **Developing probabilistic regression models:** quantile regression, density estimation.
- **Modeling forecasting residuals:** distribution fitting, residual simulation.

Introduction

PLF can be in the form of quantiles, intervals, or density functions.



Traditional PLF can only capture the probability distribution of the load individually in each period and cannot integrate dependencies among different periods.

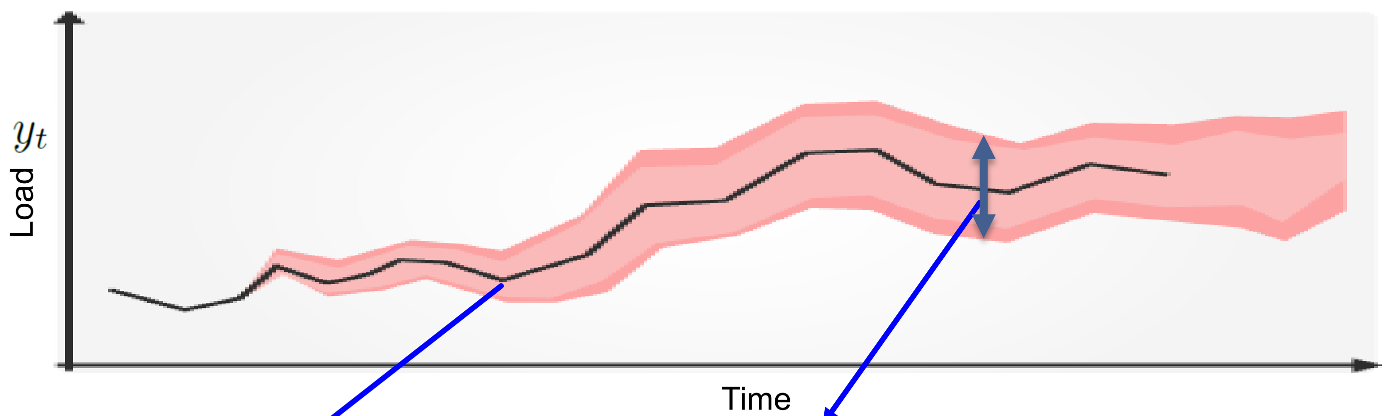
Dependencies across periods?

Generating scenarios is an effective way to capture such dependencies!

- ➢ Renewable energy scenarios generation?
- ➢ Long-term uncertainties?
- ➢ Our work focuses on **short-term load** scenarios generation.

Introduction

From the perspective of forecasting, the electrical load y_t contains two parts:



deterministic part:

$\hat{y}_t = f(\mathbf{X}_t)$

uncertain part:

$\epsilon_t = y_t - \hat{y}_t$

Train point forecast model

$f(\cdot)$

Train residual generation model

$G(\cdot)$

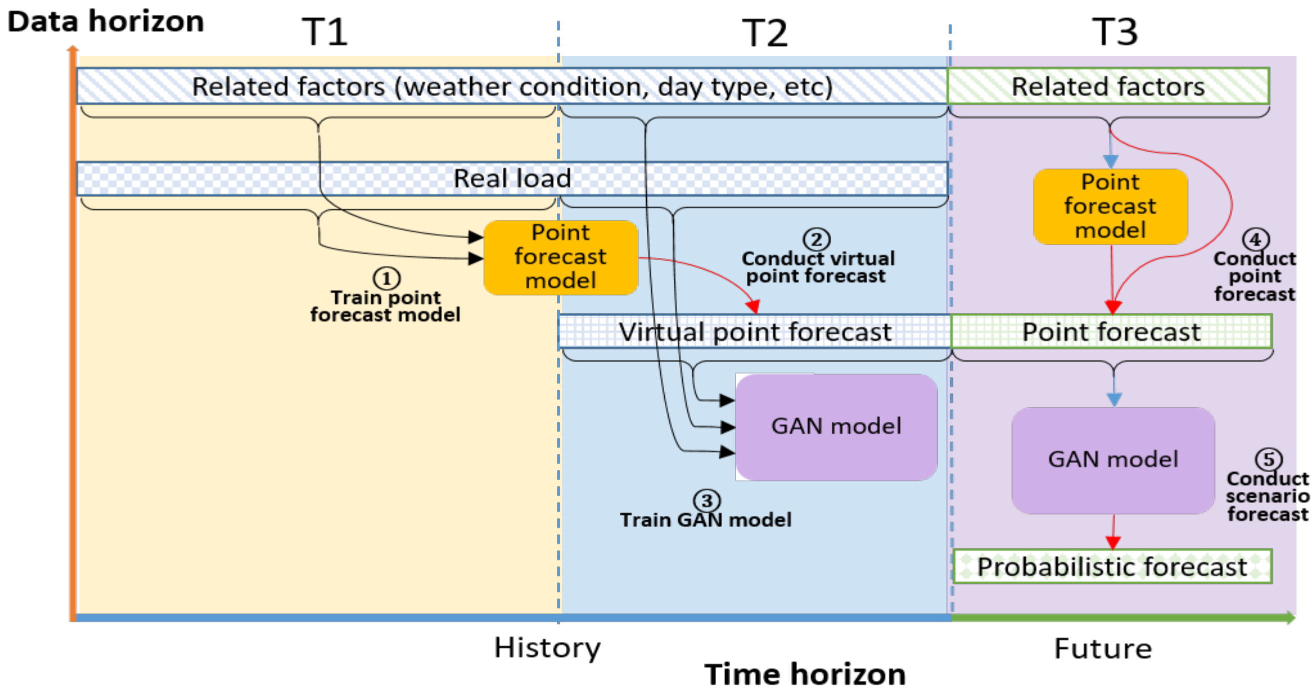
Test both models

$f(\mathbf{X}_t) + G(\mathbf{C}_t, z)$

T1	T2	T3

Methodologies

Proposed Framework

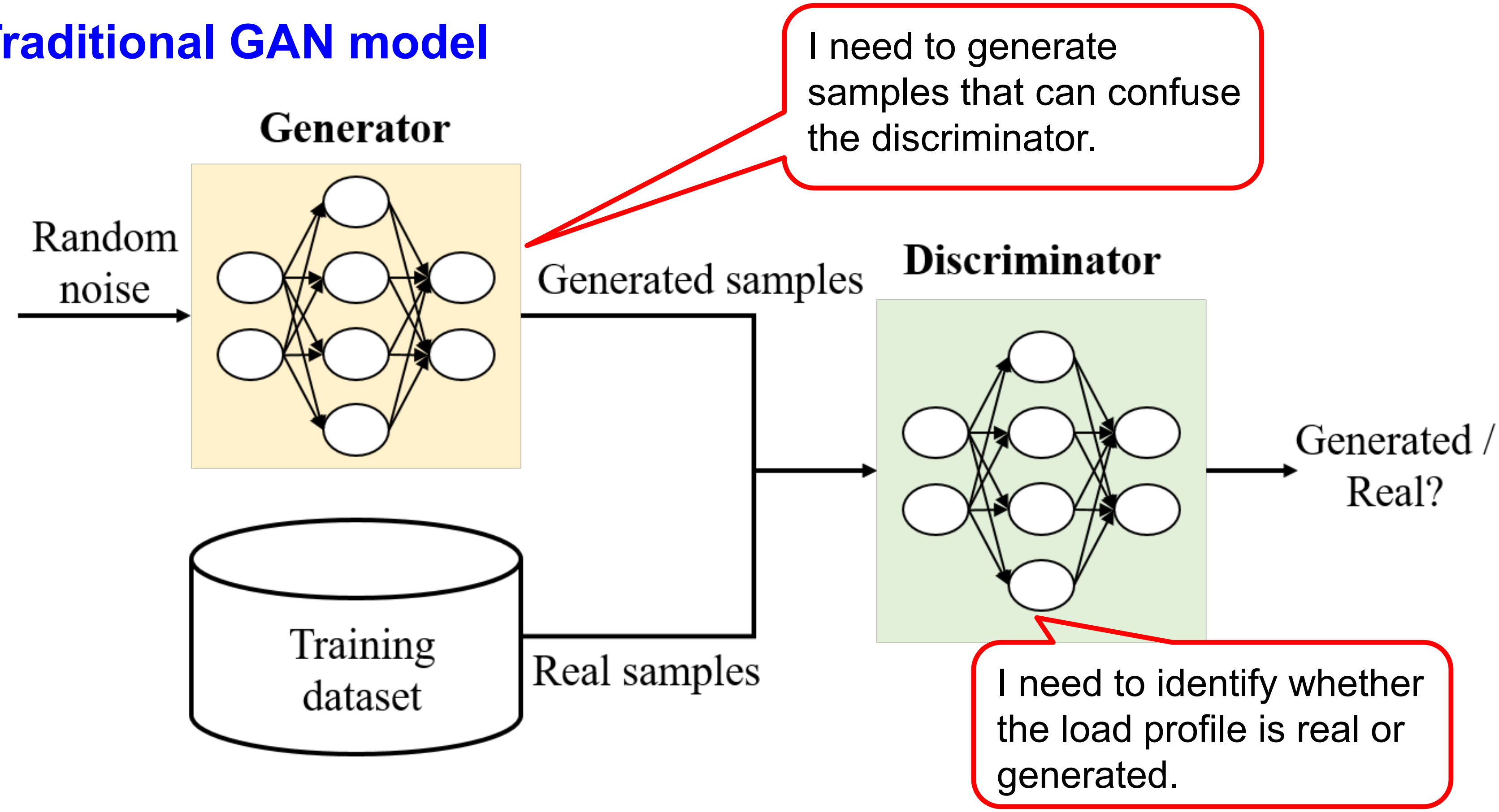


➢ **Train** $f(\mathbf{X}_t)$ using point forecasting models such as MLP/SVR/RF/GBRT	➢ **Test** $f(\mathbf{X}_t)$ ➢ **Calculate** ϵ_t ➢ **Train** $G(\mathbf{C}_t, z)$	➢ **Test** $f(\mathbf{X}_t)$ ➢ **Test** $G(\mathbf{C}_t, z)$ ➢ **Final forecasts** $f(\mathbf{X}_t) + G(\mathbf{C}_t, z)$

Formulate the generation model $G(\cdot)$ using GAN !

Methodologies

Traditional GAN model



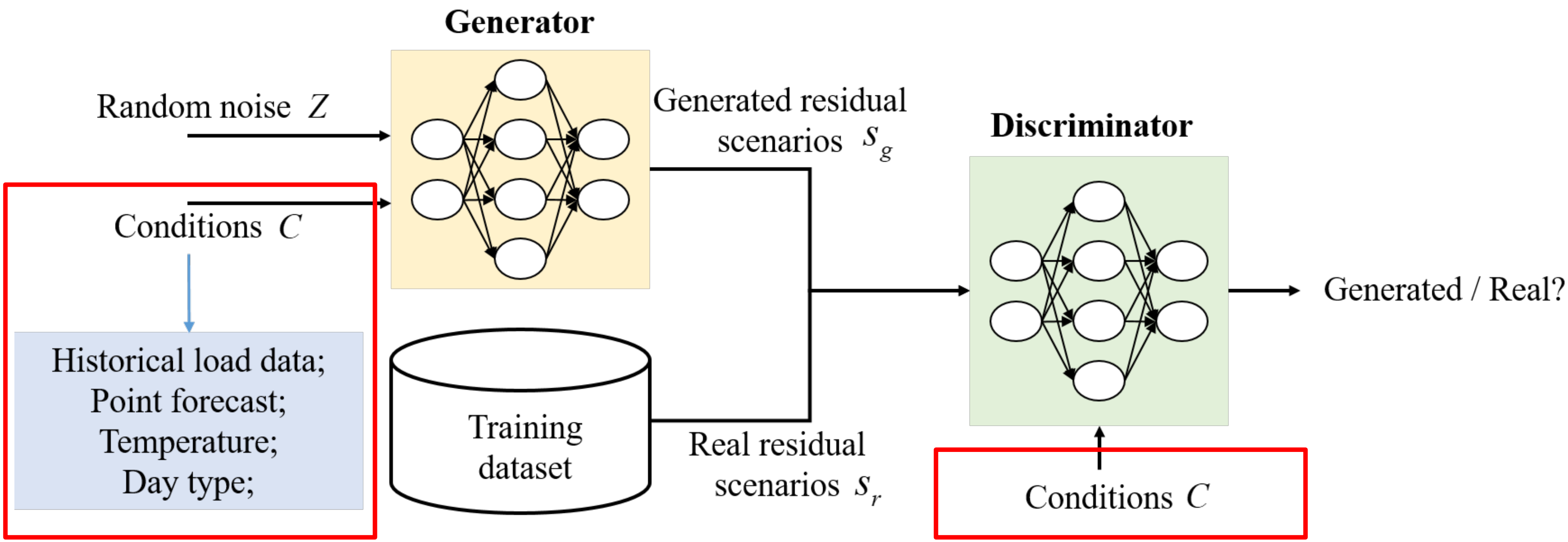
The **adversarial game** between these two neural networks can be presented as a min-max optimization model:

$$\max_{\theta_D} \mathbb{E}_{S_r}[\log(D(s_r;\theta_D))]+\mathbb{E}_Z[\log(1-D(G(z;\theta_G);\theta_D))]$$

Methodologies

Modification #1: Conditional GAN

❖ **Model the conditional relationship between residual and different factors.**

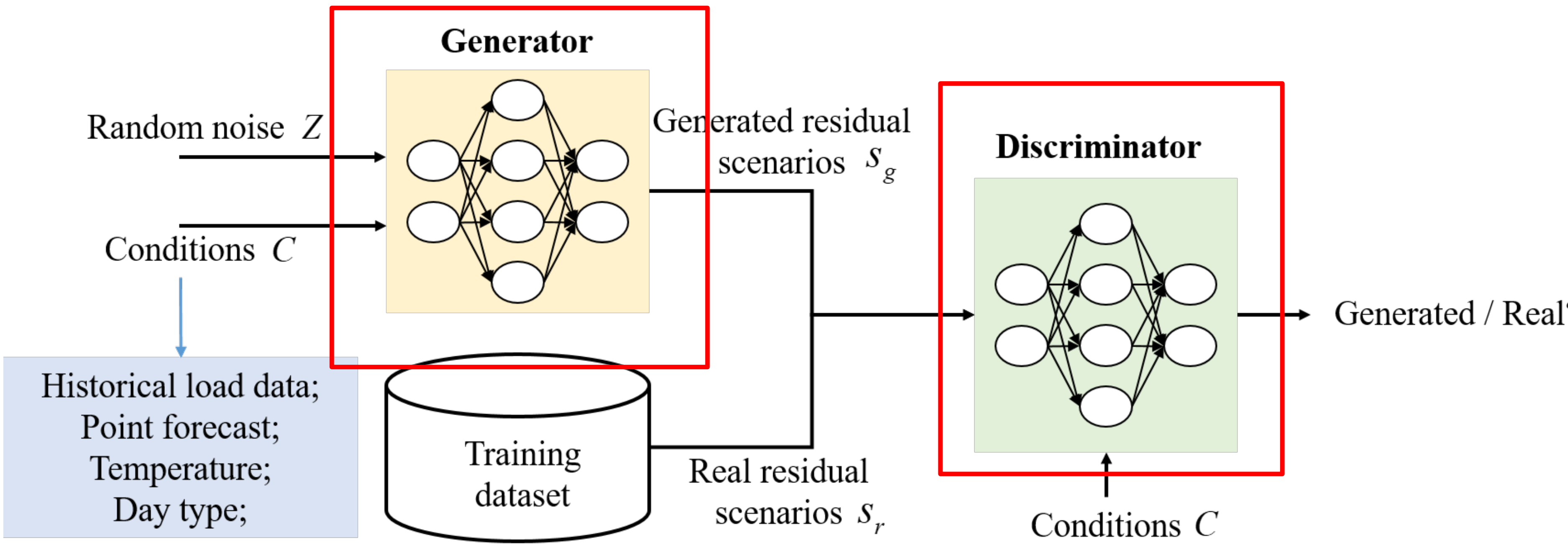


On the basis of the traditional GAN model, the conditional GAN model adds the conditions C to the inputs of the generator and discriminator.

Methodologies

Modification #2: Wasserstein GAN

- **Address the possibly unstable training process originating from the cross-entropy-based loss function.**



$$L_D = -\mathbb{E}_{S_r}[\log(D(s_r;\theta_D))] - \mathbb{E}_Z[\log(1 - D(s_g;\theta_D))]$$

$$L_G = -\mathbb{E}_Z[\log(1 - D(G(z;\theta_G);\theta_D))]$$

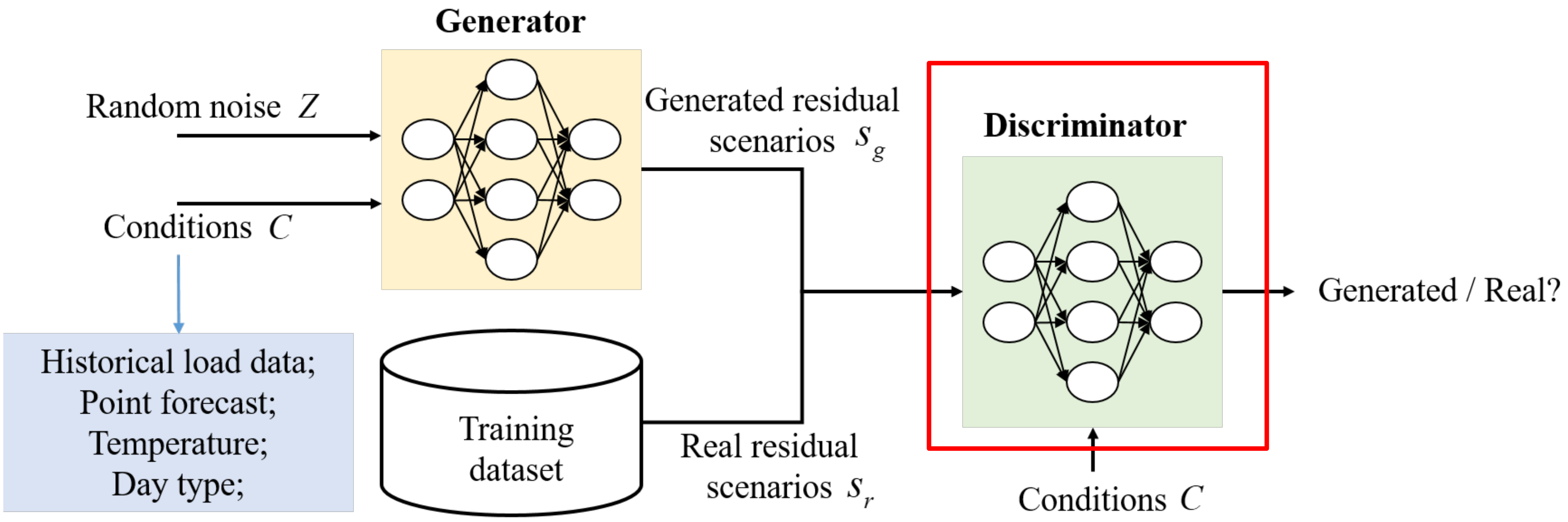
$$L_D = -\mathbb{E}_{S_r}[(D(s_r|\mathbf{C},\theta_D))] + \mathbb{E}_Z[D(s_g|\mathbf{C},\theta_D)],$$

$$L_G = -\mathbb{E}_Z[D(G(z;\theta_G)|\mathbf{C},\theta_D)],$$

Methodologies

Modification #3: Gradient Penalty

- **Reduce the risks of the gradient vanishing or exploding during the training process.**



Gradient penalty: $GP = \lambda \mathbb{E}_S[||\nabla D(\hat{s})||_2 - 1]^2$

$$L_D = -\mathbb{E}_{S_r}[(D(s_r|\mathbf{C}, \theta_D))] + \mathbb{E}_Z[D(s_g|\mathbf{C}, \theta_D)] + GP.$$

CWGAN-GP: Conditional Wasserstein GAN model with GP

Methodologies

Evaluation Criteria

From uncertainty perspective: **Pinball loss (PL)** and **Winkler Score (WS)** assess the calibration and sharpness simultaneously.

$$\mathrm{PL}\left(\hat{y}_{t,q}, y_t\right) = \begin{cases} \left(y_t - \hat{y}_{t,q}\right) q & \hat{y}_{t,q} \le y_t \\ \left(\hat{y}_{t,q} - y_t\right)(1-q) & \hat{y}_{t,q} > y_t \end{cases}$$

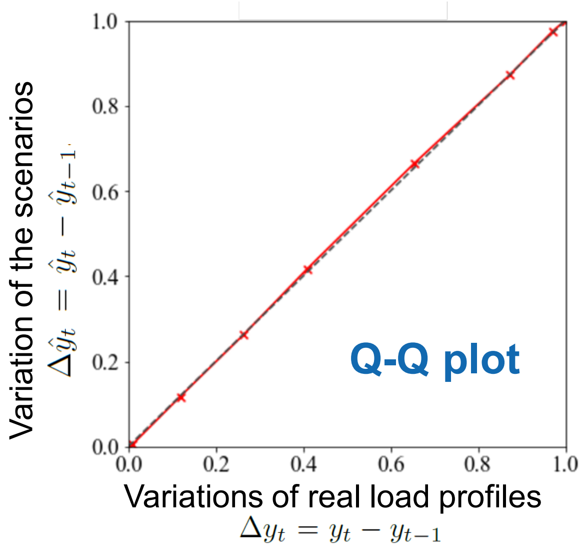
Performance of overall quantiles

$$\mathrm{WS}(L_t, U_t, y_t) = \begin{cases} \delta_t + 2(L_t - y_t)/\alpha & y_t \le L_t \\ \delta_t & L_t < y_t < U_t \\ \delta_t + 2(y_t - U_t)/\alpha & U_t \le y_t \end{cases}$$

Performance of extreme quantiles

From variation perspective:

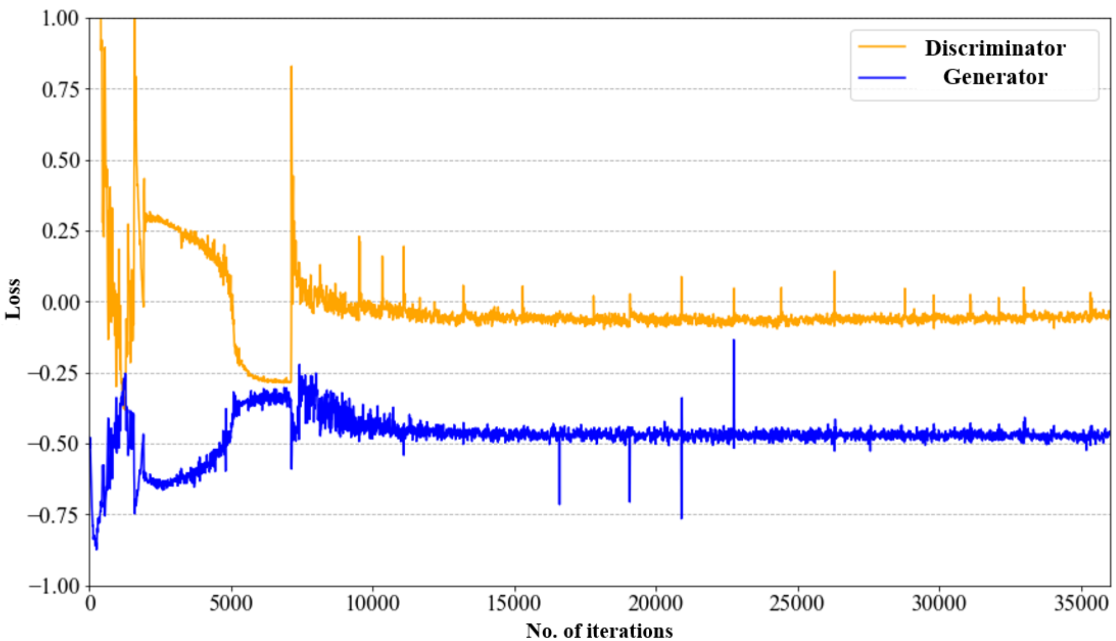
the **Q-Q plot** visually evaluate the similarity of the distributions of the variations of the real load profiles and the generated scenarios.



Case Studies

Case studies are conducted on an open load dataset from the ME area of the Independent System Operator-New England (ISO-NE).

Convergence Analysis



Losses of the generator and discriminator networks

- The losses of both generator and discriminator networks **converge without any anomaly gradient**.
- The loss function of the discriminator converges to **near zero**, indicating that the discriminator nearly fails to distinguish the real and generated samples.

Case Studies

Performance w.r.t. Uncertainty

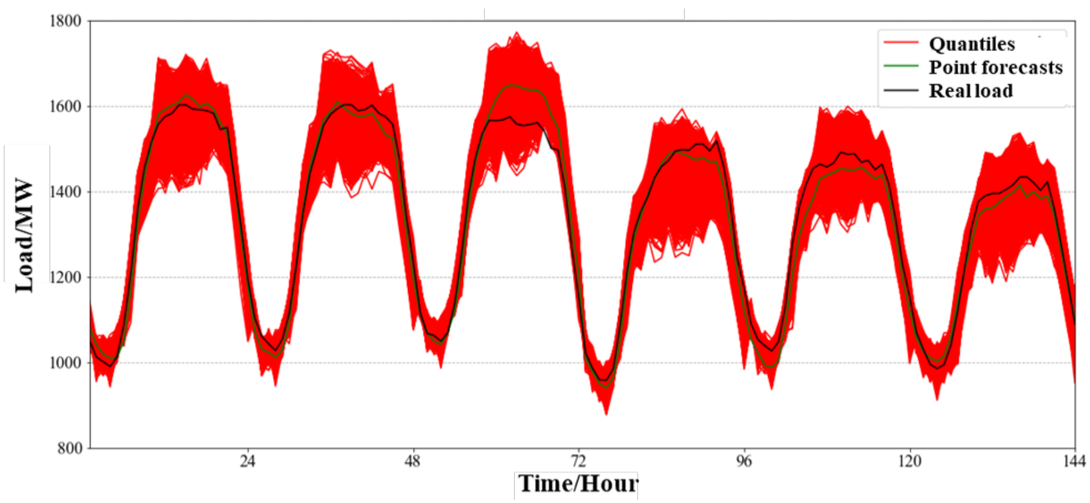
Point Forecasts	Uncertainty Modeling	PL	WS(a=0.2)	WS(a=0.1)
AVE	Proposed	**12.38**	**180.1**	262.14
	CWGAN	13.7	191.49	261.59
	QRF	14.23	189.64	**231.84**
	QGBRT	14.05	190.66	243.66
SVR	Proposed	**12.55**	**182.04**	259.84
	CWGAN	12.78	190.91	281.9
	QRF	14.5	194.66	**240.76**
	QGBRT	14.75	201.88	255.55
RF	Proposed	**12.91**	**183.32**	260.07
	CWGAN	13.1	187.58	263.7
	QRF	14.44	194.06	242.99
	QGBRT	13.94	186.66	**233.99**
GBRT	Proposed	**12.24**	**172.15**	236.34
	CWGAN	13.11	180.22	235.96
	QRF	14.06	183.18	**223.12**
	QGBRT	14.36	189.39	236.59

Comparison among different PLF methods

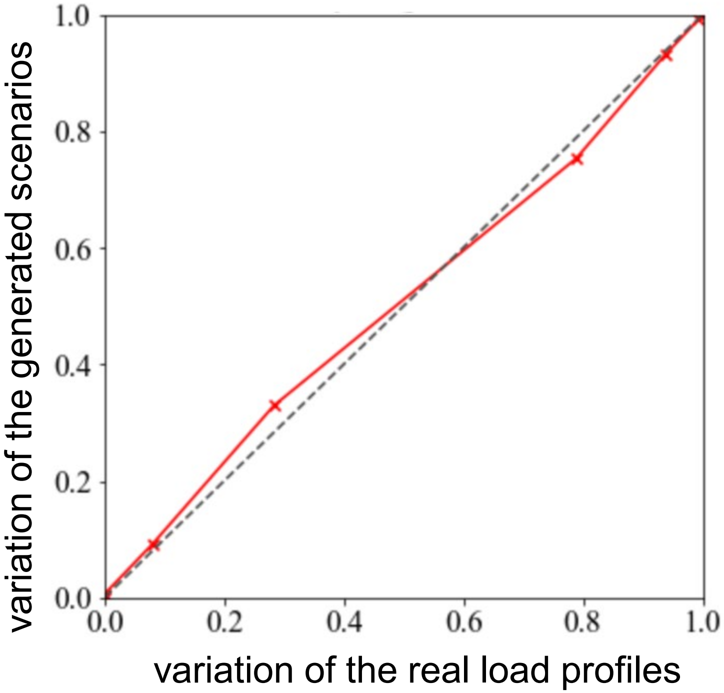
- For different point forecasts, our proposed CWGAN-GP model outperforms the CWGAN model, QRF, and QGBRT in terms of **PL and WS (a= 0.2)**.
- However, QRF instead of the CWGAN-GP model performs better in terms of **WS (a= 0.1).**

Case Studies

Performance w.r.t. Variation



10,000 generated scenarios of six selected days

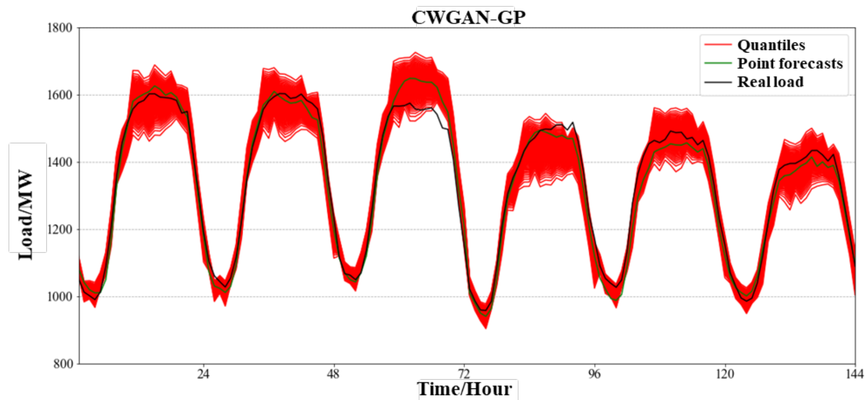


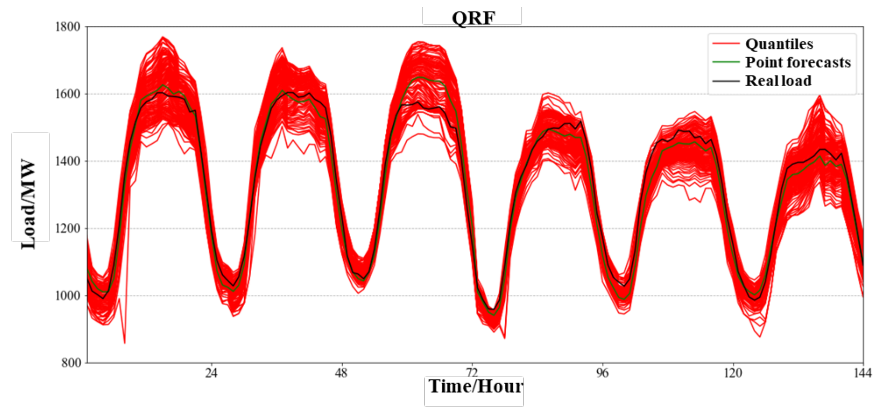
The Q-Q plot

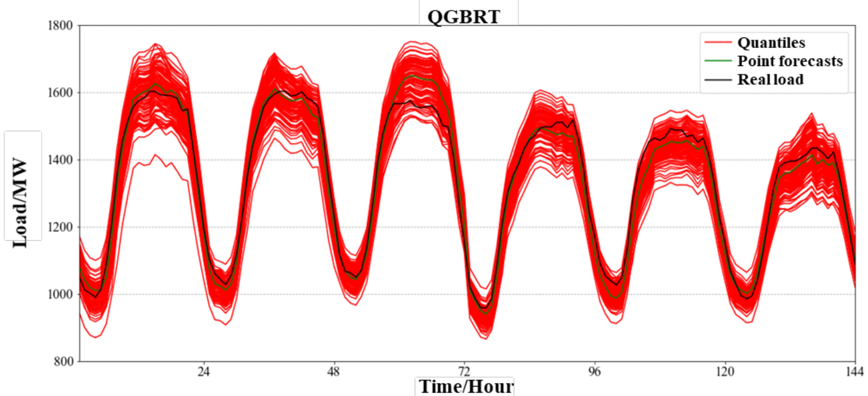
- The distribution of the variation of the generated scenarios is very similar to that of the real load profiles.
- Thus, **the generated scenarios can well represent the variations of the load profiles**.

Case Studies

Investigation on Extreme Quantiles







Quantile forecasts obtained by CWGAN-GP, QGBRT, and QRF

- The scenarios generated by the CWGAN-GP model are **more concentrated** compared with those by QGBRT and QRF.
- Such results suggest that it is **hard for the CWGAN-GP model to generate extreme scenarios.**

Conclusions

- The proposed CWGAN-GP model is capable of modeling both the uncertainties of each time period and the variations across different time periods.
- The limitation of the GAN model is that very few extreme scenarios can be generated because of the characteristics of the discriminator network.

- Future work will be focused on combining the GAN model-based forecasts and quantile regression-based forecasts to further improve the performance of the final probabilistic forecasts.

ETH zürich

Thank you for your attention

Yi Wang | yiwang@eeh.ee.ethz.ch | www.eeyiwang.com